PCT / CA 00/00738
AUGUS 2000 (09 . 08 . 00)

CA09/00738

#3

REC'D 05 SEP 2000
WIPO PCT

PA 269432

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

4

July 05, 2000

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: ***60/182,905***
FILING DATE: ***February 16, 2000***

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

H. PHILLIPS
Certifying Officer

jc714 U.S. PTO
02/16/00

PROVISIONAL APPLICATION COVER SHEET

A/PRAV

This is a request for filing a PROVISIONAL APPLICATION under 37 CFR 1.53(b)(2)

jc526 U.S. PTO
60/182905
02/16/00

Docket Number	11004-1	Type a Plus Sign (+) inside this Box -->	+

INVENTOR(S)/APPLICANT(S)

LAST NAME	FIRST NAME	Middle Initial	RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY)
JAIN	Ravinder	K.	2413 Irvine Avenue, Saskatoon, Sask. S7J 2A9, Canada
MACKENZIE	Samuel	L.	17 Cambridge Crescent, Saskatoon, Sask. S7H 3P9, Canada

TITLE OF THE INVENTION (280 CHARACTERS MAX.)

Oil Content And Composition enhancement by Targeted/Non-Targeted Expression Of Plant Plastidial And Bacterial Gene For Glycerol-3-Phosphate Acyltransferase

CORRESPONDENCE ADDRESS

J. Wayne Anderson
National Research Council of Canada
Intellectual Property Services Office, EG-10, Bldg. M-58
Montreal Road, Ottawa, Ontario, Canada K1A 0R6

STATE	Ontario	ZIP CODE	K1A 0R6	COUNTRY	Canada

ENCLOSED APPLICATION PARTS *(Check all that Apply)*

X	*Specification*	*Number of pages*	8		*Small Entity Statement*	
X	*Drawing(s)*	*Number of Sheets*	6		*Other (specify)*	

METHOD OF PAYMENT *(Check One)*

	A check or money order is enclosed to cover the Provisional filing fees	*Provisional Filing Fee Amount ($)*	$ 150.00
X	*The Commissioner is hereby authorized to charge filing fees and credit Deposit Account Number* 14-0429		

The invention was made by an agency of the United States Government or under contract with an agency of the United States Government. X No

Date:
17 February 2000

Respectfully submitted,

J. Wayne Anderson
Patent Agent for Applicant
Regn No. 28,158

Enclosurers
:ech

_____ Additional inventors are being named a separately numbered sheets attached hereto.

PROVISIONAL APPLICATION FILING ONLY

Oil Content and Composition Enhancement by Targeted/Non-Targeted Expression of Plant Plastidial and Bacterial Gene For Glycerol-3-Phosphate Acyltransferase

BACKGROUND OF THE INVENTION

There is a general industrial demand for higher oil content in living organisms in order to be more cost effective during production and processing. It is estimated that a 1% increase in seed oil would add $20 M to the value of the Canadian canola crop at the farm gate.

One way of accomplishing it is by modifying metabolic pathways of living organisms by enhancing the expression of genes involved in oil synthesis. For example, overexpression of a yeast *SLC1-1* gene has been found to increase oil content in plants (Zou et al., 1997).

Only a limited number of genes have been reported to enhance oil content (for example, Zou et al., 1997, Marillia et al., 1998). There is a need to explore other approaches using other genes which might be more effective alone or in combination with previously cloned genes.

SUMMARY OF THE INVENTION

We propose to increase seed oil content and modify seed oil composition by overexpressing the genes for the enzyme glycerol-3-phosphate acyltransferase (GPAT). This we achieved earlier in yeast (See US Provisional Patent Application No. 60/139,788, June 21, 1999). We now report the extension of our invention to plant oils.

This is the first demonstration of an increase in carbon flux into plant triglycerides by manipulating the expression of a GPAT enzyme.

Detailed Description of the Invention

GPAT is the first enzyme of membrane and storage lipid (triacylglycerol;TAG) biosynthesis in living organisms. It determines the type of fatty acid incorporated in the *sn*-1 position of membrane and storage lipids. Overall, its specificity for various acyl-CoA appears to be flexible (Frentzen, 1993; Frentzen and Wolter, 1998). Since it is the first step in triglyceride and phospholipid lipid biosynthesis, it is surmised that enhancing production of GPAT might also result in an increase of 1-acylglycerol-3-phosphate through the Kennedy pathway, resulting in higher oil content. Three types of GPAT have been reported from plants: plastidial (P), mitochondrial (M), and cytosolic (ER). They exhibit variable specificity toward acyl-CoA and acyl-ACP substrates (Frentzen, 1993). The ER form of GPAT is considered to be the most important for TAG manipulation but it has not been possible to isolate a gene for this enzyme. However, cDNAs for P-GPATs has been isolated and characterized from a number of species. Also, heterologous GPAT genes from bacteria and mammals are available (see Wilkison and Bell, 1997; Dircks and Sul, 1997).

A plastidial GAPT gene has been used by others to change the fatty acid composition of membrane lipids and to improve chilling tolerance (Murata, 1992 and Nishizawa, 1996). The bacterial GPAT gene (*plsB*) has also been used to change the fatty acid composition of membrane lipids and to decrease chilling tolerance (Wolter, 1992). None of the studies thus far have used GPAT genes to increase oil content in a living organism.

Because of the unavailability of an ER GPAT gene from plants, we set out to use a plant plastidial GPAT and a bacterial GPAT gene to establish the proof of our concept. Here we demonstrate the use of a plant plastidial or bacterial GPAT gene to increase oil content in plant seeds. Two genes were modified to create four versions to achieve higher oil content.

A cDNA (~gene) for a P-GPAT has been isolated in this lab from safflower (*Carthamus tinctorius*) (Bhella and MacKenzie, 1994). A P-GPAT can use both acyl-ACP and acyl-CoA substrates, although the latter with lesser efficiency (Frentzen, 1993). It was surmised that in the absence of any acyl-ACP in the cytosol, a P-GPAT will have to use acyl-CoA, although it has to compete with the ER-GPAT for substrates, acyl-CoAs and glycerol-3-phosphate (G-3-P). Its competitive ability could be enhanced by overexpressing it by using a strong constitutive or a tissue specific promoter. Normally, a P-GPAT is targeted to plastids by a transit peptide (TP). The TP was deleted to confine the GPAT activity in the cytosol (P-GPAT-TP). The modifed gene was labelled as *ctpgpat-tp*. It is anticipated that an ER environment may also be required for a P-GPAT-TP to effectively participate in TAG biosynthesis. The ER retention signals (ERRS; Jackson et al., 1990) were added to to P-GPAT-TP to meet this requirement. The modified gene was labelled as *ctpgpat-tp+errs*.

The *Escherichia coli's plsB* gene has been previously characterized (see Wilkison and Bell, 1997). The *plsB* can use both acyl-ACP and acyl-CoA equally effectively. The *plsB* gene was used as such and also with an added ERRS sequence. The gene constructs were labelled as *plsB* and *plsB+errs*.

EXPERIMENTAL

Preparation of chimeric genes and expression vectors

An open reading frame (*orf*; ~1.1 kb) without *tp* was amplified by PCR from the *ctpgpat* cDNA. Another chimeric gene containing an *ers* at its 3' end was also PCR amplified. The *orf* of the *E.coli plsB* gene (~2.5 kb) was PCR amplified in unmodified form or with an *ers* at its 3' end. The blunt-end PCR fragments generated using *Pfu* DNA polymerase were cloned into pSK II cloning vector. The genes were labeled as *ctpgpat-tp, ctpgpat-tp+ers, plsB* and *plsB+ers*. The intact *ctpgpat* cDNA was also used (Bhella and MacKenzie, 1994). Their sequences are shown in, respectively, Figures 1 to 5. It should be noted that the sequences representing the *ctpgpat* are different from that published earlier (Bhella and MacKenzie, 1994). There were errors in the published sequence.

For plant expression study, chimeric genes were recovered from the cloning vector and cloned into the plant transformation vector, pHS737 (tandem 35s CaMV promoter) for constitutive expression of GPATs. The recombinants were transformed into *Agrobacterium tumefaciens* GV 3101 for Arabidopsis transformation.

Transformation of *Arabidopsis thaliana*

Arabidopsis thaliana plants were transformed by the floral dip method. Seeds (T_1) from these plants were collected and selected on a growth medium containing kanamycin. Transgenic plants were grown to maturity and seeds (T_2) from 10 individual plants were collected and used for lipid analysis. Wild type and plants transformed with vector alone were grown as controls along with the transformed plants.

Lipid analysis

The fatty acid composition of the seeds was determined by GC analysis following extraction of the oil and conversion of the triglycerides to fatty acid methyl esters. A known amount of C15 triglyceride was added to the seed sample as a marker before oil extraction. Total seed lipid content was estimated on the basis of the recovery of C15 fatty acid methyl ester.

RESULTS AND DISCUSSION:

Most of the transgenic plants appeared normal in morphology. Those that were not were discarded. The oil compositions and seed sizes of selected samples representing each contruct are shown in Table 1 and Figure 5. Seeds of plants transformed using only the pHS737 vector were indistinguishable in oil content from wild type control plants grown under the same conditions. All other gene constructs produced higher seed oil content. The unmodified *spgpat*, which is expressed in the plastid, produced oil increases ranging from 10 to 21%, suggesting that *lysophosphatidic acid* is released from the plastids. On average the greatest increase in oil was observed in seeds of transformants carrying the *spgpat-tp* gene (average +22%). Addition of an ER targeting sequence had no apparent effect on seed oil content. The *plsB* gene increased seed oil content by an average of 15%. The addition of an ER targeting sequence resulted in an average seed oil increase of 18% but the difference cannot be regarded as significant. It is somewhat unexpected that the *spgpat* protein, which ordinarily acts on an acyl-ACP substrate and has a much reduced activity with acyl-CoA substrates, effected a greater seed oil increase (on average) than the *plsB* protein. The *plsB* protein reacts equally well with acyl-ACP and acyl-CoA substrates.

Seeds of plants transformed with the vector only did not differ significantly in average weight from wild type plants (Table 1). Seeds of individual plants from each contruct were significantly larger than wild type and the pHS737 control; e.g. 315-2, 301-2, 302-6, 303-3 and 304-15. However, increased seed oil content was not always positively correlated with increased seed weight; e.g. 303-7 and 304-1. Samples representing *spgpat* genes only were more highly correlated (coefficient 0.53) than those representing *plsB* genes (coefficient 0.18).

The phenotypes representing increased seed oil content and size would result in increased yields in oilseed crops. Those representing an increase in seed oil content without an increase in seed size would provide more oil per tonne of seed and would present advantages to oilseed processors. Appropriate phenotypes can, therefore, be chosen for specific applications.

We have demonstrated that expression of both a bacterial and a plant GPAT increases plant seed oil content and seed size.

Some compositional changes were also observed in the seed oils of the transformants. Selected examples are illustrated in Table 2. Seed oils from plants transformed with the vector alone (pHS737) were not significantly different from the untransformed wild type *A. thaliana*. No significant changes were observed in the proportions of any of the saturated fatty acids or of 18:2 and 20:1. However, the proportions of 18:1 Z9 (oleic acid) decreased and of 18:3 Z9, Z12, Z15 (α-linolenic acid) increased in some of the individual transformants of all constructs. Examples are 315-6, 315-8, 301-2, 302-5, 303-3 and 304-4. These compositional changes do not always coincide with an increase in total seed oil content; for example, 315-6. Furthermore, increases in seed oil content are not always accompanied by a change in composition; for example 315-4 and 303-7.

We hypothesize that overexpression of GPAT increases the proportion of lyso-phosphatidic acid which contains oleic acid at the *sn*-1 position and, therefore, in turn, the proportion of phosphatidic acid having the same *sn*-1 fatty acid composition. The latter constitutes the substrate for subsequent desaturation by the Δ-12 fatty acid desaturase to produce linoleic acid which is subsequently further desaturated by the Δ-15 fatty acid desaturase. The result is increased proportions of α-linolenic acid, the terminal product of the desaturation pathway.

It is surprising that the change in the proportions of oleic and α-linolenic acids is observed for both the safflower and *E. coli* GPATs. In *in vitro* assays, the safflower plastidial GPAT prefers unsaturated fatty acids as compared to the *E. coli* GPAT which prefers saturated fatty acids. Nevertheless, both genes, in modified and unmodified forms, present an approach to decreasing the proportion of oleic acid and increasing the proportion of linolenic acid in contexts where such a result is desirable. In the latter context, the process is independent of any intellectual property rights vested in any of the fatty acid desaturases.

Since GPAT is a ubiquitous enzyme for phospholipid and triglyceride biosynthesis and the biochemical pathway is identical in all living organisms, it is reasonable to expect that the overexpression of these genes will also enhance oil synthesis in other microorganisms and other plant species. It is also reasonable to expect that expression of other GPATs will have similar effects on seed oil content.

Summary

An increase in triglyceride synthesis and average seed size following expression in *Arabidipsis thaliana* of an *E. coli* GPAT and a modified safflower GPAT has been demonstrated.

References

Bhella RS and MacKenzie SL. 1994. Nucleotide sequence of a cDNA from *Carthamus tinctorius* encoding a glycerol-3-phosphate acyl transferase. Plant Physiol 106:1713-1714.

Dircks LK and Sul HS. 1997. Mammalian mitochondrial glycerol-3-phosphate acyltransferase. Biophys Acta 1348: 17-26.

Eccleston VS and Harwood JL. 1995. Solubilisation, partial purification and properties of acyl-CoA:glycerol-3-phosphate acyltransferase from avocado (Persea americana) fruit mesocarp. Biochimica Biophysica Acta 1257:1-10.

Frentzen M. 1993. Acyltransferases and Triacylglycerols. In Lipid Metabolism in Plants, Ed TS Moore, Jr. CRC Press, Inc., FL, pp 1195-230.

Frentzen M and Wolter FP. 1998. Molecular biology of acyltransferases involved in glycerolipid synthesis. In Plant Lipid Biosynthesis: fundamentals and agricultural applications (Ed. JL Harwood). Cambridge University Press. pp. 247-272.

Gierhart DL. 1984. Multistage process for the preparation of fats and oils. US Patent 4,485,172.

Jackson MR, Nilsson T, and Peterson PA. 1990. Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. EMBO J 9: 3153-3162.

Jamadar SC, Cao WF and Samaniego E. 1996. Relationship between adipose polyamine concentrations and triacylglycerol synthetic enzymes in lean and obese Zucker rats. Enzyme Protein 49: 222-230.

Murata N, Ishizaki-Nishizawa O, Higashi S, Hayashi H, Tasaka Y and Nishida I. 1992. Genetically engineered alteration in the chilling sensitivity of plants. Nature 356: 710-713.

Murata N. and Tasaka Y. 1997. Glycerol-3-phosphate acyltransferase in plants. Biochim Biophys Acta 1348: 10-16.

Nishizawa O. 1996. Chilling-resistant plants and their production. US patent 5,516,667.

Ratledge C. 1988. Yeasts for lipid production. Biochem Soc Trans 16: 1088-1091.

Rutar V. 1989. Magic angle sample spinning NMR spectroscopy of liquids as a nondestructive method for studies of plant seeds. J Agric Food Chem 37: 70-74.

Sinden KW. 1987. The production of lipids by fermentation within the EEC. Enzyme Microb Technol 9: 124-125.

Wilkison WO and Bell RM 1997. *sn*-Glycerol-3-phosphate acyltransferase from *Escherichia coli*. Biochim Biophys Acta 1348: 3-9.

Wolter FP, Schmidt R and Heinz E. 1992. Chilling sensitivity of *Arabidopsis thaliana* with genetically engineered membrane lipids. EMBO J 11:4685-92

Zou J, Katavic V, Giblin EM, Barton DL, MacKenzie SL, Keller WA, Hu X and Taylor DC. 1997. Modification of seed oil content and acyl composition in the Brassicaceae by expression of a yeast *sn*-2 acyltransferase gene. Plant Cell 9:909-923.

Zou J-T, Marillia E-F, Qi Q, Barton DL and Taylor DC. 1998. Does mitochondrially-generated acetate contribute to plastidial fatty acid biosynthesis ? Antisense repression of an *Arabidopsis thaliana* mitochondrial pyruvate dehydrogenase kinase (PDHK) gene and its effects on oil content and plant development. B-71. 13th International Symposium on Plant Lipids. Seville, Spain. July.

09103502.021600

Table 1. Oil content of wild type and transformant *Arabidopsis thaliana* seeds

Sample	Transformant	Oil Content (Wt %)	% Increase	Wt/100 seeds (mg)
Wild type	Average (n=7)	26.9 ± 0.8		1.42 ± 0.29
pHS737	Average (n=7)	26.8 ± 1.09		1.41 ± 0.26
ctpgpat	315-2	32.4	20.9	2.47
	315-3	29.4	9.7	1.92
	315-4	30.1	12.3	1.90
	315-7	29.6	10.4	1.85
ctpgpat - tp	301-2	34.0	26.9	2.03
	301-3	34.6	29.1	1.90
	301-5	29.3	9.3	1.79
ctpgpat-tp+errs	302-2	29.4	9.7	1.44
	302-6	32.7	22.0	2.45
plsB	303-2	29.8	11.2	1.28
	303-3	30.8	14.9	1.98
	303-4	29.1	8.6	1.89
	303-7	33.2	23.9	1.53
plsB + errs	304-1	30.9	15.2	1.51
	304-2	32.5	21.3	1.60
	304-3	30.3	13.1	1.38
	304-15	32.5	21.3	2.20

Table 2. Fatty Acid Composition of Wild Type and Transformed *Arabidopsis thaliana*

Sample/construct	Wt % Oil	18:1	18:3
Wild type	26.9 ± 0.8 (n=7)	14.4 ± 0.7	15.7 ± 0.8
pHS737	26.8 ± 1.1 (n=6)	12.4 ± 1.3	16.9 ± 0.4
spgpat			
315-2	32.4	13.6	18.4
315-4	30.1	14.7	16.5
315-6	25.2	9.1	18.4
315-8	27.1	9.2	19.9
315-10	28.3	12.4	17.5
spgpat-tp			
301-1	27.5	12.0	17.1
301-2	34.0	10.9	18.8
301-3	34.6	12.2	18.2
spgpat-tp+errs			
302-4	28.1	12.5	16.5
302-5	25.6	12.2	17.5
302-6	32.7	12.9	17.5
302-7	27.8	15.8	16.5
plsb			
303-1	28.4	12.0	17.8
303-3	30.8	11.4	17.6
303-4	29.1	14.7	17.1
303-7	33.2	14.8	15.9
plsb + errs			
304-2	32.5	12.4	17.0
304-3	30.3	12.1	16.8
304-4	23.5	10.7	18.0
304-9	27.4	13.9	17.4
304-15	32.5	14.1	16.9

Figure 1. Nucleotide sequence of the gene *ctpgpat*(*Carthamus tinctorius* plastidial) used for transforming yeast and *Arabidopsis thaliana*. The start and stop codons are underlined. Nucleotides that are different from the published sequence are highlighted.

TCTCTCTCTCTCACACACAACACACAAAACACACACTACTGCTACTTTCTCTCTCTACTACACTCTCCTCTCGCTATG
TCGATCTTCTTCTCTCCTTCCTCCCCTACTCTCTTCTTCTCCACCACAAACGCAAATCCTAGGGTTTCTCCTTCATCT
TCACCTTCTTCTGCCTTCACTCCTCCTCTGTCTTCTTCTCGCCTCCGCCCGATTCTCCGGGGGGTTTCCGTGCCTCGCG
TTCTCTGCGCCGGCGAATGCCGCCCATGGCACGGCGGAGACCGTCCACGGCAACAAGTGGCCGTCACCGTCGTCCTCC
TCCTCTGCTGCTACGCAACCGTCCGCTGGATCCGACCACGGTCACTCTCGTACATTCATCGATGCTCGTTCCGAACAA
GATCTTCTTTCTGGAATTCAAAGAGAGTTGGAAGCTGGAACACTGCCAAAACATATTGCTCAAGCAATGGAGGAGCIA
TATCAGAACTACAAAAATGCAGTTCTCCAAAGTGCGGCTCCTCATGCAGAAGATATTGTGTTGTCAAACATGCGTGTA
GCGTTTGATCGIATGTTCTTGGATGTGAAGGAGCCGTTTGAATTTTCACCATATCATGAAGCTATTTTGGAACCTTTT
AACTACTATATGTTTGGTCAAAATTATATTCGGCCTTTGGTCAATTTCAGGGAATCATACGTTGGCAATGTCTCCGTT
TTCGGTGTAATGGAAGAGCAGCTTAAGCAGGGTGACAACGTGGTTTTGATCTCAAACCATCAAACAGAAGCAGATCCA
GCTGTTATTGCCTTGATGCTTGAAACAACAAACCCCCATATTTCTGAGAACATAATCTACGTGGCAGGGGATAGAGTA
ATAACAGATCCTCTTTGCAAGCCTTTCAGCATGGGAAGGAATCTGTTGTGCGTGTATTCAAAAAGCATATGAATGAT
GTTCCTGAGCTTGCTGAGATGAAAAAAAGATCAAATACAAGAAGTTTAAAAGAGGTGGCTTTGCTTTTGAGGGGCGGA
TCTAAAATAATATGGATTGCGCCAAGTGGTGGCAGGGACAGGCCAGATCCTATCACAAATCAGTGGTTTCCGGCACCG
TTTGATGCCACTTCGCTTGACAACATGAGAAGGCTCGTGGACCATGCTGGTTTGGTGGGTCACATATATCCTTTAGCC
ATATTGTGCCATGACATCATGCCCCCTCCTCTTCAGGTTGAGAAAGAAATTGGAGAGAAGAGGAGATCTCTTTTCAT
GGCACCGGAATATCAGTGGCACCGGAAATTAATTTCCAAGAAGTTACTGCCTCTTGTGGGTCCCCCCGAGGAGGCGAAG
GCAGCTTATTCACAGGCACTCTATGATTCCGTGTGTGAACAATACAAGGTGCTACATTCTGCGGTACATGGAGGAAAA
GGGTTAGAAGCATCAACACCAAGTGTCTCGTTGTCACAACCCTTGCAGTTTCTCGATTAGTCTCTTGGTTTAGAGGAG
GTGAAAGCATATTCTTTTGTTTAGATGACATAGGTGTATAGATGATACCGAAGAATAGATGTACAAACAAGTGATAGA
AAGATGTATGTCTAATCAAAAAATGTTTTCTGCATCTTGTAAAGGGATCTTCAAAACAGGCCTTTTATTTTAGCTGCA
GCAACCAATATATCAAAACAGGTTTTTCTTTTT

The translated protein sequence of the gene *ctpgpat* used for transforming yeast and *Arabidopsis thaliana*. Amino acids that are different from the published sequence are highlighted.

SLSLTHNTQNTHYCYFLSLLHSPLAMSIFFSPSSPTLFFSTTNANPRVSPSSSPSSAFTPPLSSSRLRPILRGFPCLA
FSAPANAAHGTAETVHGNKWPSPSSSSSAATQPSAGSDEGHSRTFIDARSEQDLLSGIQRELEAGTLPKHIAQAMEEL
YQNYKNAVLQSAAPHAEDIVLSNMRVAFDRMFLDVKEPFEFSPYHEAILEPFNYYMFGQNYIRPLVNFRESYVGNVSV
FGVMEEQLKQGDNVVLISNHQTEADPAVIALMLETTNPHISENIIYVAGDRVITDPLCKPFSMGRNLLCVYSKKHMND
VPELAEMKKRSNTRSLKEMALLLRGGSKIIWIAPSGGRDRPDPITNQWFPAPFDATSLQNMRRLVDHAGLVGHIYPLA
ILCHDIMPPPLQVEKEIGEKEISFHGTGISVAPEINFQEVTASCGSPEEAKAAYSQALYDSVCEQYKVLHSAVHGGK
GLEASTPSVSLSQPLQPLD

Figure 2. Nucleotide sequence of the gene *ctpgpat-tp* (*Carthamus tinctorius* plastidial GPAT from which the sequence for a putative transit peptide was removed) used for transforming yeast and *Arabidopsis thaliana*. The start and stop codons are underlined. ATG was introduced as a start codon in place of GAC. Nucleotides that are different from the published sequence are highlighted.

GGATCCATGCACGGTCACTCTCGTACATTCATCGATGCTCGTTCCGAACAAGATCTTCTTTCTGGAATTCAAAGAGAG
TTGGAAGCTGGAACACTGCCAAAACATATTGCTCAAGCAATGGAGGAGCTATATCAGAACTACAAAAATGCAGTTCTC
CAAAGTGCGGCTCCTCATGCAGAAGATATTGTGTTGTCAAACATGCGTGTAGCGTTTGATCGTATGTTCTTGGATGTG
AAGGAGCCGTTTGAATTTCACCATATCATGAAGCTATTTTGGAACCTTTTAACTACTATATGTTTGGTCAAAATTAT
ATTCGGCCTTTGGTCAATTTCAGGGAATCATACGTTGGCAATGTCTCCGTTTTCGGTGTAATGGAAGAGCAGCTTAAG
CAGGGTGACAAGGTGGTTTTGATCTCAAACCATCAAACAGAAGCAGATCCAGCTGTTATTGCCTTGATGCTTGAAACA
ACAAACCCCCATATTTCTGAGAACATAATCTACGTGGCAGGGGATAGAGTAATAACAGATCCTCTTTGCAAGCCTTTC
AGCATGGGAAGGAATCTGTTGTGCGTGTATTCAAAAAAGCATATGAATGATGTTCCTGAGCTTGCTGAGATGAAAAAA
AGATCAAATACAAGAAGTTTAAAAGGGATGGCTTTGCTTTTGAGGGCGGATCTAAAATAATATGGATTGCGCCAAGT
GGTGGCAGGGACAGGCCAGATCCTATCACAAATCAGTGGTTTCCGGCACCGTTTGATGCCACTTCGCTTGACAACATG
AGAAGGCTCGTGGACCATGCTGGTTTGGTGGGTCACATATATCCTTTAGCCATATTGTGCCATGACATCATGCCCCCT
CCTCTTCAGGTTGAGAAAGAAATTGGAGAGAAGAGGGATCTCTTTTCATGGCACCGGAATATCAGTGGCACCGGAA
ATTAATTTCCAAGAAGTTACTGCCTCTTGTGGGTCCCCCGAGGAGGCGAAGGCAGCTTATTCACAGGCACTCTATGAT
TCCGTGTGTGAACAATACAAGGTGCTACATTCTGCGGTACATGGAGGAAAAGGGTTAGAAGCATCAACACCAAGTGTC
TCGTTGTCACAACCCTTGCAGTTTCTCGATTAGGATCC

The translated protein sequence of the gene *ctpgpat-tp* used for transforming yeast and *Arabidopsis thaliana*. T he mature polypeptide will not have the first 90 amino acids of the *ctpgpat* (Bhella and MacKenzie, 1994). Amino acid 91, Asp (D), is replaced with Met (M). Amino acids that are different from the published sequence are highlighted.

GSMRGHSRTFIDARSEQDLLSGIQRELEAGTLPKKIAQAMEELYQNYKNAVLQSAAPHAEDIVLSNMRVAFDRMFLDV
KEPFEFSPYREAILEPFNYYMFGQNYIRPLVNFRESYVGNVSVFGVMEEQLKQGDKVVLISNHQTEADPAVIALMLET
TNPRISENIIYVAGDRVITDPLCKPFSMGRNLLCVYSKKHMNDVPELAEMKKRSNTRSLKEMALLLRGGSKIIWIAPS
GGRDRPDPITNQWFPAPFDAISLDNMRRLVDHAGLVGHIYPLAILCHDIMPPPLQVEKEIGEKEISFHGTGISVAPE
INFQEVTASCGSPEEAKAAYSQALYDSVCEQYKVLHSAVHGGKGLEASTPSVSLSQPLQFLD.D

Figure 3. Nucleotide sequence of the gene *ctpgpat-tp+errs* (*Carthamus tinctorius* plastidial GPAT from which the sequence for a putative transit peptide was removed and sequences for endoplasmic reticulum retention signals were added). Start and stop codons are underlined. ATG was introduced as a start codon in place of GAC. Additionally, AAG and AAA (also underlined) in place of TTG and TTT were introduced for ER retention of the protein. Nucleotides that are different from the published sequence are highlighted.

GGATCCATGCACGGTCACTCTCGTACATTCATCGATGCTCGTTCCGAACAAGATCTTCTTTCTGGAATTCAAAGAGAG
TTGGAAGCTGGAACACTGCCAAAACATATTGCTCAAGCAATGGAGGAGCTATATCAGAACTACAAAAATGCAGTTCTC
CAAAGTGCGGCTCCTCATGCAGAAGATATTGTGTTGTCAAACATGCGTGTAGCGTTTGATCGTATGTTCTTGGATGTG
AAGGAGCCGTTTGAATTTTCACCATATCATGAAGCTATTTTGGAACCTTTTAACTACTATATGTTTGGTCAAAATTAT
ATTCGGCCTTTGGTCAATTTCAGGGAATCATACGTTGGCAATGTCTCCGTTTTCGGTGTAATGGAAGAGCAGCTTAAG
CAGGGTGACAAGGTGGTTTTGATCTCAAACCATCAAACAGAAGCAGATCCAGCTGTTATTGCCTTGATGCTTGAAACA
ACAAACCCCCATATTTCTGAGAACATAATCTACGTGGCAGGGGATAGAGTAATAACAGATCCTCTTTGCAAGCCTTTC
AGCATGGGAAGGAATCTGTTGTGCGTGTATTCAAAAAAGCATATGAATGATGTTCCTGAGCTTGCTGAGATGAAAAAA
AGATCAAATACAAGAAGTTTAAAAG[illegible]TGGCTTTGCTTTTGAGGGGCGGATCTAAAATAATATGGATTGCGCCAAGT
GGTGGCAGGGACAGGCCAGATCCTATCACAAATCAGTGGTTTCCGGCACCGTTTGATGCCACTTCGCTTGACAACATG
AGAAGGCTCGTGGACCATGCTGGTTTGGTGGGTCACATATATCCTTTAGCCATATTGTGCCATGACATCATGCCCCCT
CCTCTTCAGGTTGAGAAAGAAATTGGAGAGAAGAG[illegible]GATCTCTTTTCATGGCACCGGAATATCAGTGGCACCGGAA
ATTAATTTCCAAGAAGTTACTGCCTCTTGTGGGTCCCCCGAGGAGGCGAAGGCAGCTTATTCACAGGCACTCTATGAT
TCCGTGTGTGAACAATACAAGGTGCTACATTCTGCGGTACATGGAGGAAAAGGGTTAGAAGCATCAACACCAAGTGTC
TCGTTGTCACAACCCAAGCAGAAACTCGATTAGGATCC

Translated protein sequence of the gene *ctpgpat-tp+errs*. The mature polypeptide will not have the first 90 amino acids of the *ctgpat* (Bhella and MacKenzie, 1994). Amino acid 91, Asp (D), is replaced with Met (M). Also Lys (K), underlined, replaced a Phe (F) and Leu (L) at position 3 and 5 from the C terminus. Amino acids that are different from the published sequence are highlighted.

GSMHGHSRTFIDARSEQDLLSGIQRELEAGTLPKHIAQAMEELYQNYKNAVLQSAAPHAEDIVLSNMRVAFDRMFLDV
KEPFEFSPYHEAILEPFNYYMFGQNYIRPLVNFRESYVGNVSVFGVMEEQLKQGDKVVLISNHQTEADPAVIALMLET
TNPHISENIIYVAGDRVITDPLCKPFSMGRNLLCVYSKKHMNDVPELAEMKKRSNTRSLK[illegible]MALLLRGGSKIIWIAPS
GGRDRPDPITNQWFPAPFDATSLDNMRRLVDHAGLVGHIYPLAILCHDIMPPPLQVEKEIGEK[illegible]ISFHGTGISVAPE
INFQEVTASCGSPEEAKAAYSQALYDSVCEQYKVLHSAVHGGKGLEASTPSVSLSQPKQKLD.D

Figure 4. Nucleotide sequence of the gene *plsB* (*E.coli* GPAT). Start and stop codons are underlined. A nucleotide that is different from the published sequence is highlighted. It was introduced during PCR.

AGATCTTCCCATGACTTTCTGCTATCCTTGCCGCGCATTTGCATTATTAACCAGAGGCTTTACATCGTTTATGTCCGG
CTGGCCACGAATTTACTACAAATTACTGAATTTACCATTAAGCATCCTGGTAAAAAGCAAGTCTATTCCGGCAGATCC
TGCCCCGGAACTGGGGCTGGATACCTCTCGTCCAATTATGTACGTTTACCGTACAACTCGAAAGCAGATTTGCTGAC
GTTGCGCGCCCAGTGTCTGGCACATGACTTGCCTGACCCGTTAGAGCCGCTGGAAATCGACGGCACGCTACTGCCGCG
CTATGTGTTCATTCACGGCGGGCCGCGTGTGTTCACCTATTACACGCCGAAAGAAGAGTCTATTAAGCTGTTCCACGA
CTATCTCGATTTGCACCGTAGCAACCCAAATCTGGATGTGCAGATGGTGCCAGTGTCGGTGATGTTTGGTCGCGCGCC
GGGGCGTGAAAAAGGCGAAGTGAACCCGCCGCTGCGTATGCTTAACGGCGTACAGAAATTTTCGCTGTACTGTGGCT
CGGTCGCGACAGTTTTGTGCGTTTCTCGCCGTCAGTTTCGCTGCGCCGTATGGCGGATGAACACGGCACGGATAAAAC
TATCGCTCAGAAACTGGCGCGCGTGGCGCGTATGCACTTTGCCCGTCAACGTCTGGCTGCCGTAGGCCCACGTTTGCC
TGCTCGTCAGGATCTGTTTAATAAGCTGCTCGCCTCCCGCGCCATTGCCAAAGCGGTAGAAGATGAAGCGCGCAGCAA
AAAAATCTCCCATGAAAAAGCGCAGCAGAACGCGATTGCACTGATGGAAGAGATTGCGGCGAATTTCTCTTACGAGAT
GATTCGCCTGACTGACCGTATTCTGGGCTTCACCTGGAACCGACTTTACCAGGGCATCAACGTCCATAACGCTGAGCG
CGTTCGCCAGCTGGCCCACGACGGTCATGAGCTGGTATATGTGCCTTGCCACCGCAGTCACATGGACTACCTGCTGCT
TTCTTACGTGCTGTATCACCAGGGGCTGGTGCCGCCGCATATCGGCGCCGGGATCAACCTGAATTTCTGGCCTGCCGG
GCCGATTTTCCGCCGTCTGGGGGCGTTCTTTATTCGCCGTACGTTTAAAGGCAATAAACTTTATTCCACCGTTTTCCG
GGAGTATCTCGGCGAACTGTTCAGCCGTGGTTATTCCGTCGAGTACTTCGTGGAAGGCGGTCGTTCCCGTACGGGGCG
TTTGCTGGATCCGAAAACTGGTACGCTGTCGATGACCATTCAGGCGATGCTGCGTGGCGGCACGCGTCCGATTACGCT
GATTCCGATCTATATCGGTTATGAGCACGTCATGGAAGTGGGTACTTACGCCAAAGAACTGCGCGGCGCGACGAAAGA
GAAAGAGAGCCTGCCGCAGATGCTGCGCGGTTTAAGCAAGCTGCGTAATCTCGGTCAGGGTTACGTCAACTTCGGTGA
ACCAATGCCGTTGATGACCTACCTTAACCAGCATGTACCTGACTGGCGTGAATCTATCGATCCCATCGAAGCGGTGCG
TCCGGCATGGTTAACGCCGACGGTCAATAATATTGCTGCCGATCTGATGGTACGCATTAACAACGCAGGCGCGGCAAA
CGCCATGAACCTGTGCTGTACTGCGCTACTGGCATCACGTCAGCGCTCACTCACCCGCGAGCAGTTAACCGAGCAACT
CAACTGCTACCTGGATCTGATGCGCAACGTGCCCTACTCCACGGACTCTACCGTTCCTTCAGCCAGCGCCAGCGAGCT
TATCGATCACGCGCTGCAAATGAACAAGTTTGAAGTCGAGAAAGACACAATCGGCGACATCATCATTCTGCCGCGCGA
GCAAGCGGTGCTGATGACCTACTATCGCAACAACATTGCGCATATGTTGGTGCTGCCTTCGCTGATGGCGTCAATCGT
CACCCAGCATCGCCACATCTCCCGCGACGTATTGATGGAGCACGTCAATGTGCTTTACCCAATGCTGAAAGCGGAGCT
GTTCCTGCGCTGGGATCGCGACGAGTTGCCGGACGTTATTGATGCGCTGGCAAATGAGATGCAACGTCAGGGGCTGAT
TACCCTGCAAGATGATGAGTTGCATATCAACCCGGCGCATTCTCGCACGCTACAGCTGCTGGCCGCAGGCGCGCGCGA
AACGCTGCAACGTTATGCCATCACCTTCTGGTTGTTGAGTGCCAACCCGTCGATCAACCGCGGTACGCTGGAGAAAGA
GAGCCGCACCGTCGCGCAACGTCTCTCCGTGCTGCACGGCATCAACGCGCCGGAGTTCTTCGACAAGGCGGTGTTCAG
TTCTCTGGTGCTGACACTGCGTGATGAAGGGTATATCAGCGATAGCGGCGATGCCGAACCGGCAGAAACGATGAAGGT
TTATCAGTTGCTGGCGGAGTTGATTACATCAGACGTGCGTTTGACGATTGAGAGTGCGACGCAGGGCGAAGGGTAATC
AGATCT

Translated protein sequence of the gene *plsB* (*E.coli* GPAT). An amino acid that is different from the published sequence is highlighted.

MTFCYPCRAFALLTRGFTSFMSGWPRIYYKLLNLPLSILVKSKSIPADPAPELGLDTNRPIMYVLPYNSKADLLTLRA
QCLAHDLPDPLEPLEIDGTLLPRYVFIHGGPRVFTYYTPKEESIKLFHDYLDLHRSNFNLDVQMVPVSVMFGRAPGRE
KGEVNPPLRMLNGVQKFFAVLWLGRDSFVRFSPSVSLRRMADEHGTDKTIAQKLARVARMHFARQRLAAVGPRLPARQ
DLFNKLLASRAIAKAVEDEARSKKISHEKAQQNAIALMEEIAANFSYEMIRLTDRILGFTWNRLYQGINVHNAERVRQ
LAHDGHELVYVPCHRSHMDYLLLSYVLYHQGLVPPHIAAGINLNFWPAGPIFRRLGAFFIRRTFKGNKLYSTVFREYL
GELPSRGYSVEYFVEGGRSRTGRLLDPKTGTLSMTIQAMLRGGTRPITLIPIYIGYEHVMEVGTYAKELRGATKEKES
LPQMLRGLSKLRNLGQGYVNFGEPMPLMTYLNQHVPDWRESIDPIEAVRPAWLTPTVNNIAADLMVRINNAGAANAMN
LCCTALLASRQRSLTREQLTEQLNCYLDLMRNVPYSTDSTVPSASASELIDHALQMKKFEVEKDTIGDIIILPREQAV
LMTYYRNNIARMLVLPSLMAIIVTQHRHISRDVLMERVNVLYPMLKAELFLRWDRDFLPDVIDALANEMQRQGLITLQ
DDELHINPAHSRTLQLLAAGARETLQRYAITFWLLSANPSINRGTLEKESRTVAQRLSVLHGINAPEFFDKAVFSSLV
LTLRDEGYISDSGDAEPAETMKVYQLLAELITSDVRLTIESATQGEG.SD

Figure 5. Nucleotide sequence of the gene *plsB+errs* (*E.coli* GPAT to which sequences for endoplasmic reticulum retention signals were added). Start and stop codons are underlined. AAG and AAG (also underlined) in place of GCC and ACG were introduced at the 3' end for ER retention of protein.

AGATCTTCCCATGACTTTCTGCTATCCTTGCCGCCGCATTTGCATTATTAACCAGAGGCTTTACATCGTTTATGTCCGG
CTGGCCACGAATTTACTACAAATTACTGAATTTACCATTAAGCATCCTGGTAAAAAGCAAGTCTATTCCGGCAGATCC
TGCCCCGGAACTGGGGCTGGATACCTCTCGTCCAATTATGTACGTTTTACCGTACAACTCGAAAGCAGATTTGCTGAC
GTTGCGCGCCCAGTGTCTGGCACATGACTTGCCTGACCCGTTAGAGCCGCTGGAAATCGACGGCACGCTACTGCCGCG
CTATGTGTTCATTCACGGCGGGCCGCGTGTGTTCACCTATTACACGCCGAAAGAAGAGTCTATTAAGCTGTTCCACGA
CTATCTCGATTTGCACCGTAGCAACCCAAATCTGGATGTGCAGATGGTGCCAGTGTCGGTGATGTTTGGTCGCGCGCC
GGGGCGTGAAAAAGGCGAAGTGAACCCGCCGCTGCGTATGCTTAACGGCGTACAGAAATTTTCGCTGTACTGTGGCT
CGGTCGCGACAGTTTTGTGCGTTTCTCGCCGTCAGTTTCGCTGCGCCGTATGGCGGATGAACACGGCACGGATAAAAC
TATCGCTCAGAAACTGGCGCGCGTGGCGCGTATGCACTTTGCCCGTCAACGTCTGGCTGCCGTAGGCCCACGTTTGCC
TGCTCGTCAGGATCTGTTTAATAAGCTGCTCGCCTCCCGCGCCATTGCCAAAGCGGTAGAAGATGAAGCGCGCAGCAA
AAAAATCTCCCATGAAAAAGCGCAGCAGAACGCGATTGCACTGATGGAAGAGATTGCGGCGAATTTCTCTTACGAGAT
GATTCGCCTGACTGACCGTATTCTGGGCTTCACCTGGAACCGACTTTACCAGGGCATCAACGTCCATAACGCTGAGCG
CGTTCGCCAGCTGGCCCACGACGGTCATGAGCTGGTATATGTGCCTTGCCACCGCAGTCGACATGGACTACCTGCTGCT
TTCTTACGTGCTGTATCACCAGGGGCTGGTGCCGCCGCATATCGCCGCCGGGATCAACCCGAATTTCTGGCCTGCCGG
GCCGATTTTCCGCCGTCTGGGGGCGTTCTTTATTCGCCGTACGTTTAAAGGCAATAAACTTTATTCCACCGTTTCCG
GGAGTATCTCGGCGAACTGTTCAGCCGTGGTTATTCCGTCGAGTACTTCGTGGAAGGCGGTCGTTCCCGTACGGGGCG
TTTGCTGGATCCGAAAACTGGTACGCTGTCGATGACCATTCAGGCGATGCTGCGTGGCGGCACGCGTCCGATTACGCT
GATTCCGATCTATATCGGTTATGAGCACGTCATGGAAGTGGGTACTTACGCCAAAGAACTGCGCGGCGCGACGAAAGA
GAAAGAGAGCCTGCCGCAGATGCTGCGCGGTTTAAGCAAGCTGCGTAATCTCGGTCAGGGTTACGTCAACTTCGGTGA
ACCAATGCCGTTGATGACCTACCTTAACCAGCATGTACCTGACTGGCGTGAATCTATCGATCCCATCGAAGCGGTGCG
TCCGGCATGGTTAACGCCGACGGTCAATAATATTGCTGCCGATCTGATGGTACGCATTAACAACGCAGGCGCGGCAAA
CGCCATGAACCTGTGCTGTACTGCGCTACTGGCATCACGTCAGCGCTCACTCACCCGCGAGCAGTTAACCGAGCAACT
CAACTGCTACCTGGATCTGATGCGCAACGTGCCCTACTCCACGGACTCTACCGTTCCTTCAGCCAGCGCCAGCGAGCT
TATCGATCACGCGCTGCAAATGAACAAGTTTGAAGTCGAGAAAGACACAATCGGCGACATCATCATTCTGCCGCGCGA
GCAAGCGGTGCTGATGACCTACTATCGCAACAACATTGCGCATATGTTGGTGCTGCCTTCGCTGATGGCGGCAATCGT
CACCCAGCATCGCCACATCTCCCGCGACGTATTGATGGAGCACGTCAATGTGCTTTACCCAATGCTGAAAGCGGAGCT
GTTCCTGCGCTGGGATCGCGACGAGTTGCCGGACGTTATTGATGCGCTGGCAAATGAGATGCAACGTCAGGGGCTGAT
TACCCTGCAAGATGATGAGTTGCATATCAACCCGGCGCATTCTCGCACGCTACAGCTGCTGGCCGCAGGCGCGCGCGA
AACGCTGCAACGTTATGCCATCACCTTCTGGTTGTTGAGTGCCAACCCGTCGATCAACCGCGGTACGCTGGAGAAAGA
GAGCCGCACCGTCGCGCAACGTCTCTCCGTGCTGCACGGCATCAACGCGCCGGAGTTCTTCGACAAGGCGGTGTTCAG
TTCTCTGGTGCTGACACTGCGTGATGAAGGGTATATCAGCGATAGCGGCGATGCCGAACCGGCAGAAACGATGAAGGT
TTATCAGTTGCTGGCGGAGTTGATTACATCAGACGTGCGTTTGACGATTGAGAGTGCGAAGCAGAAGGAAGGGTAATC
AGATCT

Translated protein sequence of the gene *plsB+errs*. Amino acid Lys (K), underlined, replaced a Gly (G) and Thr (T) at position 3 and 5 from the C terminus.

MTFCYPCRAFALLTRGFTSFMSGWPRIYYKLINLPLSILVKSKSIPADPAPELGLDTSRPIMYVLPYNSKADLLTLRA
QCLAHDLPDPLEPLEIDGTLLPRYVFIHGGPRVFTYYTPKEESIKLFHDYLDLHRSNPNLDVQMVPVSVMFGRAPGRE
KGEVNPPLRMLNGVQKFFAVLWLGRDSFVRFSPSVSLRRMADEHGTDKTLAQKLARVARMHFARQRLAAVGPRLPARQ
DLFNKLLASRAIAKAVEDEARSKKISHEKAQQNAIALMEEIAANFSYEMIRLTDRILGFTWNRLYQGINVHNAERVRQ
LAHDGHELVYVPCHRSHMDYLLLSYVLYHQGLVPPHIAAGINLNFWPAGPIFRRLGAFFIRRTFKGNKLYSTVFREYL
GELFSRGYSVEYFVEGGRSRTGRLLDPKTGTLSMTIQAMLRGGTRPITLIPIYIGYEHVMEVGTYAKELRGATKEKES
LPQMLRGLSKLRNLGQGYVNFGEPMPLMTYLNQHVPDWRESIDPIEAVRPAWLTPTVNNIAADIMVRINNAGAANAMN
LCCTALLASRQRSLTREQLTEQLNCYLDIMRNVPYSTDSTVPSASASELIDHALQHNKFEVEKDTIGDIIILPREQAV
LMTYYRNNIARMLVLPSLMAAIVTQHRHISRDVLMEHVNVLYPMLKAELFLRWDRDELPDVIDALANEMQRQGLITLQ
DDELHINPAESRTLQLLAAGARETLQRYAITFWLLSANPSINRGTLEKESRTVAQVLSVLHGINAPEFFDKAVTSSLV
LTLRDEGYISDSGDAEPAETMKVYQLLAELITSDVRLTIESAKQKEG

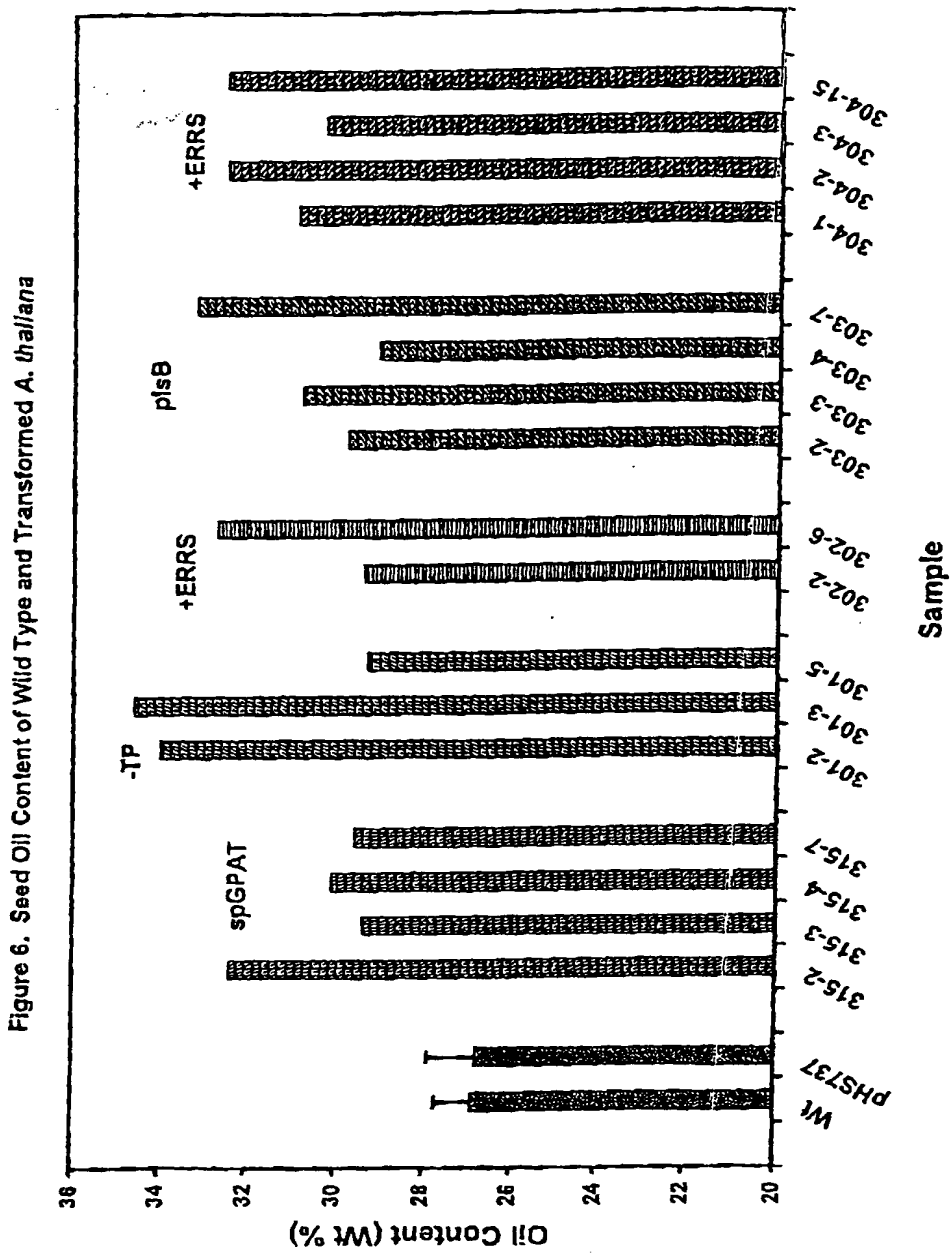

Figure 6. Seed Oil Content of Wild Type and Transformed *A. thaliana*